



Classification of pressure ulcer tissues with 3D convolutional neural network

Begoña García-Zapirain¹ · Mohammed Elmogy^{2,3} · Ayman El-Baz³ · Adel S. Elmaghreby⁴

Received: 23 October 2017 / Accepted: 19 April 2018 / Published online: 15 June 2018
© International Federation for Medical and Biological Engineering 2018

Abstract

A 3D convolution neural network (CNN) of deep learning architecture is supplied with essential visual features to accurately classify and segment granulation, necrotic eschar, and slough tissues in pressure ulcer color images. After finding a region of interest (ROI), the features are extracted from both the original and convolved with a pre-selected Gaussian kernel 3D HSI images, combined with first-order models of current and prior visual appearance. The models approximate empirical marginal probability distributions of voxel-wise signals with linear combinations of discrete Gaussians (LCDG). The framework was trained and tested on 193 color pressure ulcer images. The classification accuracy and robustness were evaluated using the Dice similarity coefficient (DSC), the percentage area distance (PAD), and the area under the ROC curve (AUC). The obtained preliminary DSC of 92%, PAD of 13%, and AUC of 95% are promising.

Keywords Pressure ulcer · 3D convolution neural network (CNN) · Tissue classification · Linear combinations of discrete Gaussians (LCDG)

1 Introduction

Chronic wounds (CWs) are considered a significant threat to the public health and economy. They have a harmful effect

The first and second authors contributed equally to this work and shared the first authorship.

The third and fourth authors are sharing the senior authorship.

✉ Ayman El-Baz
aselba01@louisville.edu

Begoña García-Zapirain
mbgarciazapi@deusto.es

Mohammed Elmogy
melmogy@mans.edu.eg

Adel S. Elmaghreby
adel.elmaghreby@louisville.edu

¹ Facultad Ingeniería, Universidad de Deusto,
Avda/Universidades 24, 48007 Bilbao, Spain

² Information Technology Department, Faculty of Computers and Information, Mansoura University, Mansoura 35516, Egypt

³ Bioengineering Department, University of Louisville, Louisville, KY 40292, USA

⁴ Department of Computer Engineering and Computer Science, University of Louisville, Louisville, KY 40292, USA

on patients causing social separation, depression, and high treatment costs [1]. In the USA, there are approximately 6.5 million patients are influenced by chronic wounds. More than US\$25 billion is spent yearly on the treatment of these wounds, because of increasing health care costs, elder people, and the incidence of diabetes and obesity [2]. CW can be defined as the wound that is remaining unhealed for longer than six weeks. Pressure ulcer (PU), diabetic foot ulcers, and venous ulcers are considered as the essential types of CW. In addition, malignant ulcer, burns, and Pyoderma gangrenosum are still quite significant types of CW. These previously mentioned types are often resistant to healing and required long-term medical care [3].

A PU or ischemic wound is developed due to lingering pressure on the part of the body causing a lack of blood flow to the tissues. In particular, when the body is not moving enough the circulatory system provides less oxygen resulting in an ischemic process under the patient's skin [4]. Up to 95% of PUs are said to be avoidable, but it is not easy to diagnose the PU until the wound becomes visible due to its advanced state [5]. In 2013, they caused 29,000 deaths and affected up from 2 to 28% of nursing home residents over the world [6], health deterioration usually occurring due to infection when detection and diagnosis are delayed [7].

PUs, also called pressure sores, bed sores [8], and pressure injuries, are frequent at support body points coinciding with prominences or bone ridges. The most susceptible sites are sacrum, heel, external and external malleoli, buttocks, trochanters, opplated, isquion, occiput, elbows, iliac crest, ears, spiny apophysis, inner sides and outer faces of knees, and side feet edges. Preventive measures reduce compression and local occlusions of blood capillaries by a shorter stay without movement. Hospital stay times differ widely, and prevalence rates of the pressure ulcers range from 12 to 16% in general acute care, 4.3 to 32% in long-term care, and 2.9 to 19% in home care [9]. The annual mean prevalence rate decreased from 7.8% in 2005 to 1.4% in 2011, being kept to 2014 [10].

This paper focuses on automated detection of PUs of grades 3 and 4. The grade 3 tissue loss forms a deep crater (sometimes covered with necrotic tissue) through the skin, even reaching the deep dermis and hypodermis. The grade 4 injuries are caves or sinuous routes due to the total skin loss with frequent destruction or tissue necrosis to muscle, bones or support structures like sinus capsule tendon.

To find the ulcer, one has to describe and locate the lesion, which can be characterized by extension, depth, and color of the tissues. Then, the skin and surrounding tissues are classified by color (pigmented, pale, cyanosis, rosy), texture (rough, thick, thin), turgidity (good, bad), temperature, humidity (dry, wet, normal), edema's degree, and location. This paper considers only on the tissue classification problem. Figure 1 demonstrates main (granulated, necrotic eschar, and slough) types of PU tissues.

In literature, PU image analysis is an active field of research. Many studies are conducted to segment and classify different tissue types to help dermatologists evaluating and diagnosing the progress of PU. For example, Dorileo et al. [11] implemented a PU segmentation system that was based on analyzing both RGB and HIS color spaces of

the PU images. They segmented each tissue type depending on choosing a threshold for various color components from the two tested color spaces. They tested the proposed system on 172 PU images and achieved accuracy of $61 \pm 25\%$.

Vereda et al. [12] proposed an automatic tissue classification system for wound images based on artificial neural networks (ANNs) and Bayesian classifiers. They used the mean shift and region-growing techniques to segment the area of the ulcer. The color and texture features are supplied to a set of k multilayer perceptrons and Bayesian classifier to classify different types of tissues in the ulcer image. They tested their system on 113 PU images and achieved accuracy of 91.5%.

Azevedo-Marques et al. [13] proposed a clustering segmentation method to segment different ulcer tissues. Their method was based on the color components in hue-saturation histograms and mathematical morphology. The proposed system was tested on 172 ulcer images. It achieved an average Jaccard coefficient of $56 \pm 22\%$ between the resulting segmentation and the ground truth (GT) that was generated by a dermatologist.

Mukherjee et al. [1] developed a tissue classification system for CW. First, they transformed the original RGB images to HIS color space to provide higher contrast. The CW images were segmented by using fuzzy divergence based thresholding to minimize edge ambiguity. They classified different CW tissues by using support vector machine (SVM) with 3rd-order polynomial kernel based on the extracted color and textural features. Their system achieved an overall accuracy of 88%.

Ahmed Fauzi et al. [2] implemented a segmentation technique for CW color images. First, they generated a red-yellow-black-white (RYKW) probability map to guide the segmentation techniques. The red, yellow, and black probability maps are utilized to handle the granulation,

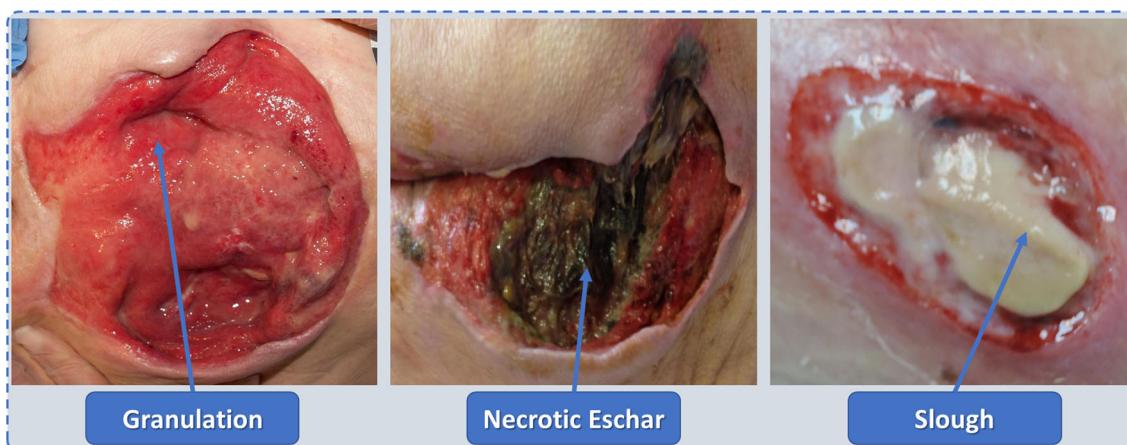


Fig. 1 Examples of the most common types of tissues in PU images

slough, and eschar wound tissues. On the other hand, the white probability map is generated to detect the white label card that was used for measurement calibration aims. Then, the region growing or optimal thresholding is used to segment tissues of the chronic wound. Finally, they ran the experiments on 80 images and achieved an average accuracy of 75.1%.

Vereda et al. [14] proposed a tissue segmentation system for PU images based on k-means clustering technique. They used ANN, random forest (RF), and SVM classifiers to distinguish different ulcer tissues. To improve the efficacy of the classifier and decrease the number of necessary predictors, they selected features based on a wrapper approach with recursive feature elimination. RF and SVM classifiers achieved the highest performance values based on extracted features. Their system run on 113 PU images and achieved an average accuracy of 88%.

Ortiz et al. [15] proposed a PU segmentation system that is based on toroidal geometry to extract various contrast levels from the proposed images. They used Otsu's threshold technique to segment the images. Then, morphological operators are used to refine the final result. They achieved average accuracy of 89% on 51 PU images.

On the other hand, some studies used deep learning to segment PU images. For instance, Wang et al. [16] proposed a wound segmentation technique based on convolution neural network (CNN) model where ConvNet features are used in infection detection via SVM classifier and in healing prediction process via Gaussian Process (GP) Regression. The results showed that ConvNet had a better accuracy compared to SVM classifier. Kawahara and Hamarneh [17] built a skin lesion segmentation system using multi-track CNN, which extended pertained CNNs for multi-resolution skin lesion classification. The results showed a higher classification accuracy compared to multi-scale approaches. Esteva et al. [18] classified the skin cancer classification by using Google's Inception v3 CNN architecture pre-trained on the 1000 object classes. The final classification layer is removed and replaced by their data set. The results showed good accuracy for the classification with a level of competencies comparable to dermatologists.

The main limitations of the current research studies can be summarized in the following points. First, there is a need to improve the accuracy of the automatic classification for both external and internal ulcer boundaries. Most of the current work determined the region of interest (ROI) manually. Second, the objective parameters associated with the morphology of the PU need to be optimally calculated. Finally, the dermatologists need a real-time automatic segmentation system for PU assessment.

To overcome the first and third limitations, we proposed an automated segmentation system to segment and classify different tissues from PU colored images. The proposed

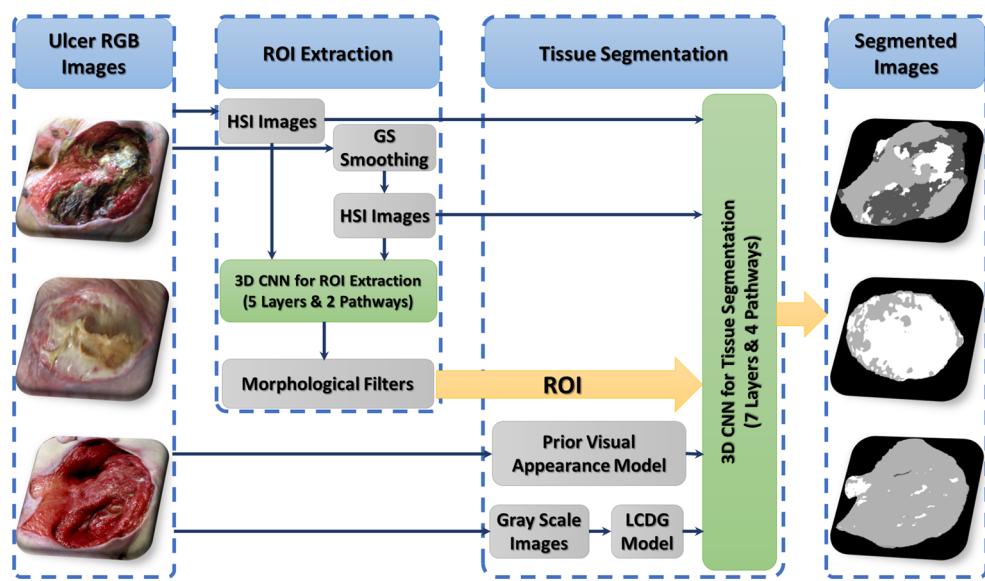
system is based on extracting and fusing different features from ulcer images to provide an accurate ulcer segmentation framework. The extracted feature modalities are provided to a 3D CNN network to segment ulcer tissues depending on various features. The main contributions of the proposed framework can be summarized in the following points. First, we proposed an automatic PU segmentation framework to segment both external and internal ulcer boundaries. Second, the proposed system is based on extracting four different feature types, which are the original image in HSI color space, the convolved image by using 3D Gaussian kernel, the prior appearance model, and the current appearance model using our linear combinations of discrete Gaussians (LCDG). Third, these modalities are fused and supplied to 3D CNN networks with multiple paths to extract the ROI and different ulcer tissues. Finally, we evaluated the performance of the proposed framework by using different metrics to show its effectiveness. The rest of this paper is organized as follows. Section 2 presents the methods that are applied during the presented research to construct the proposed system. Section 3 describes the experimental results. The results are discussed in Section 4. Finally, Section 5 summarizes the conclusion of our work and our future directions of research.

2 Methods

This work aims to develop an automatic segmentation system to detect and segment PU RGB-colored images. The main idea of the proposed system depends on extracting different features and representing them as various modalities of PU images. These modalities are fused and supplied to 3D CNN network with multiple paths to extract the ROI and different ulcer tissues. The main architecture of the proposed system is shown in Fig. 2. The proposed system consists of two main stages, which are the ROI extraction and tissue segmentation. Figure 3 illustrates the structure of the 3D CNN network that is used in the ROI extraction stage. It automatically extracts the ROI from the processed image to distinguish the PU area from other patient's skin and background objects. Two different modalities of the processed images are supplied to the 3D CNN network, which are the HSI images and the convolved HSI image with the 3D Gaussian kernel.

In addition, Fig. 4 illustrates the structure of the second 3D CNN network that is used in tissue segmentation stage. It segments the resulting ROI to extract different ulcer tissues, which are granulation (red), necrotic eschar (black), and slough (yellow) tissues. Four different modalities of the processed images are supplied to a 3D CNN network. These modalities are the above two HSI modalities in addition to the output of our current appearance model

Fig. 2 The proposed PU segmentation system



that is generated by applying a LCDG technique and the prior visual appearance model. The details of the proposed system are discussed in the following subsections. Before discussing the main stages of the proposed system, a list of the mathematical notations, which are utilized throughout the paper, is given below.

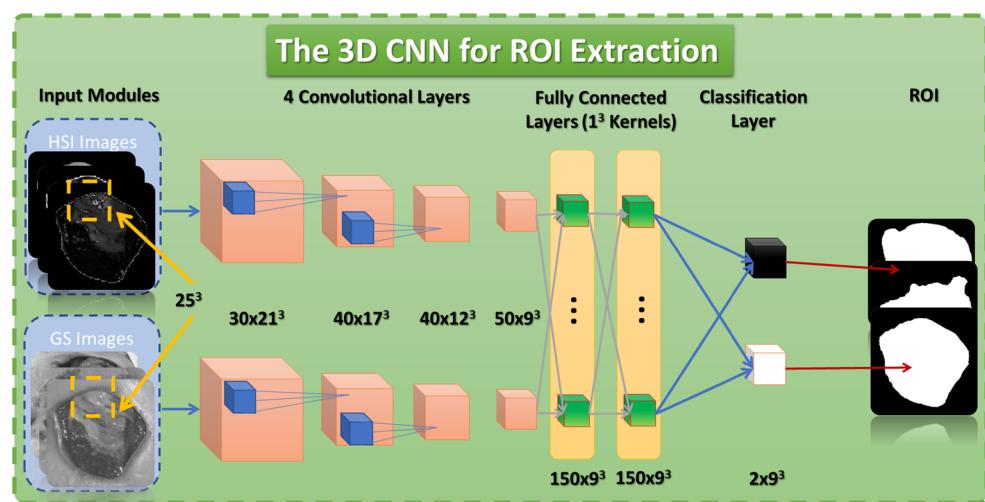
Basic notation Below $\mathbb{R} = \{\mathbf{r} = (x, y, z) : 0 \leq x \leq X - 1, 0 \leq y \leq Y - 1, 0 \leq z \leq Z - 1\}$ will denote a 3D finite arithmetic lattice of $X \times Y \times Z$ voxels, \mathbf{r} . It supports an RGB color PU image with integer intensity levels, $q \in \mathbb{Q} = \{0, 1, \dots, Q - 1\}$ in each color channel, $\mathbf{g} = \{g_r : \mathbf{r} \in \mathbb{R}, g_r \in \mathbb{Q}\}$. Binary labels, $l \in \mathbb{L}_{\text{ROI}} = \{0, 1\}$, will indicate an extracted ROI (0 for background and 1 for ROI voxels) in a ROI map, $\mathbf{m}_{\text{ROI}} = \{m_{\mathbf{r}} : \mathbf{r} \in \mathbb{R}, m_{\mathbf{r}} \in \mathbb{L}_{\text{ROI}}\}$, and four color-coded labels, $\mathbb{L}_{\text{SEG}} = \{0, 1, 2, 3\}$, will specify a segmentation map, $\mathbf{m}_{\text{SEG}} = \{m_{\mathbf{r}} : \mathbf{r} \in \mathbb{R}, m_{\mathbf{r}} \in \mathbb{L}_{\text{SEG}}\}$.

Here, the segmentation labels 0, 1, 2, and 3 are for background, necrotic eschar (black), granulation (red), and slough (white & yellow), respectively.

2.1 The ROI extraction stage

The goal of this stage is to extract the ROI, which contains the PU area, and build its map, \mathbf{m}_{ROI} . This stage uses the initial and convolved ulcer RGB images and converts them into the HSI color images to be supplied as inputs for 3D CNN pathways. The HSI color model eliminates impacts of illumination changes onto chrominance of the images. In addition, the saturation (S) channel of the HSI model provides higher contrast [1]. Gaussian smoothing (GS) of the initial RGB image is performed by convolution of the image with a moving 3D Gaussian kernel (h_{σ}). In order to add longer range properties to the original intensities

Fig. 3 The proposed 3D CNN network for the ROI extraction stage



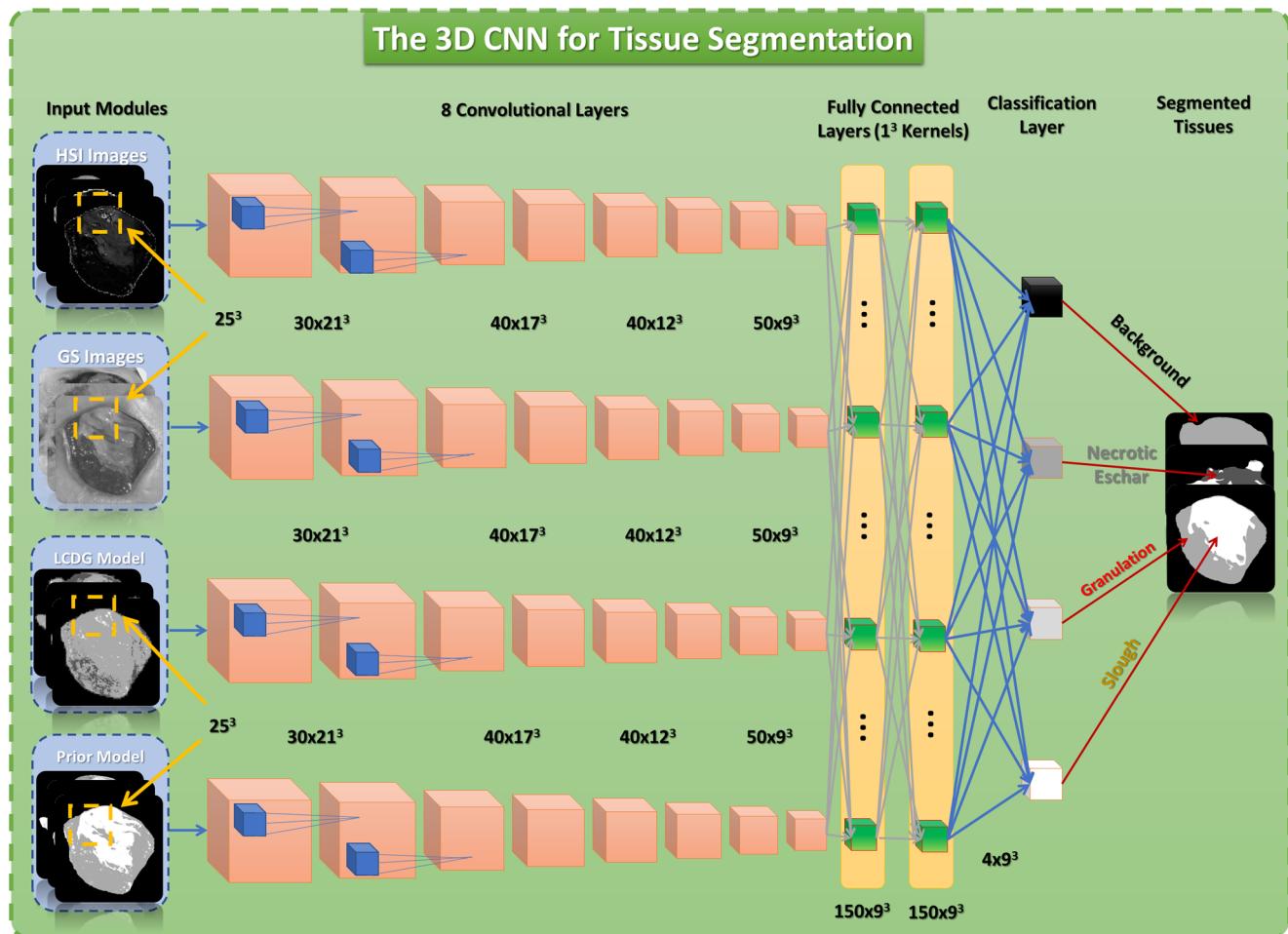


Fig. 4 The proposed 3D CNN network for the tissue segmentation stage

and their pairwise co-occurrences in the nearest $5 \times 5 \times 3$ -neighborhoods of each voxel:

$$\tilde{\mathbf{g}}_j = \mathbf{g}_j * \mathbf{h}_\sigma; \quad j \in \{R, G, B\} \quad (1)$$

where $\tilde{\mathbf{g}}_j$ represents the color channel, j , of the convolved input RGB image, \mathbf{g} , and $\mathbf{h}_\sigma = (h_{\sigma:\mathbf{r}-\mathbf{r}_c} : (\mathbf{r}, \mathbf{r}_c) \in \mathbb{R}^2)$ is a 3D Gaussian kernel with a fixed standard deviation σ [19, 20];

$$h_{\sigma:\mathbf{r}-\mathbf{r}_c} = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} |\mathbf{r} - \mathbf{r}_c|^2\right) \quad (2)$$

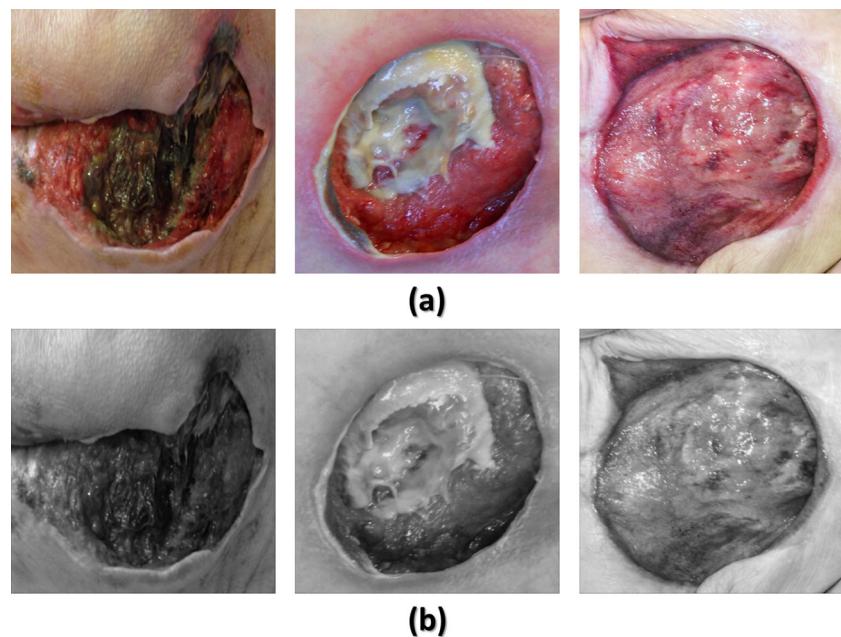
Here, $|\mathbf{r} - \mathbf{r}_c| = \sqrt{(x - x_c)^2 + (y - y_c)^2 + (z - z_c)^2}$ is the Cartesian distance between the voxel, \mathbf{r} , and the kernel's center, $\mathbf{r}_c = (x_c, y_c, z_c)$.

Both the initial RGB image and its convolved version are then converted to the HSI images, \mathbf{g}_{HSI} and \mathbf{g}_{GS} , respectively. The Gaussian smoothing helps to reduce noise in the processed images. In particular, it integrates longer-range image properties and eliminates intensity distortions that can affect quality of the captured image. However, it flattens the intensity (I) channel by decreasing its maximal and increasing its minimal values. The image is convolved

by the $(5 \times 5 \times 3)$ kernel with $\sigma^2 = 2.5$ to preserve more details. Figure 5 shows examples of ulcer images with their convolved versions.

The mentioned two models are supplied as separate inputs to a 3D CNN network, which is called DeepMedic [21]. The DeepMedic network is a 3D CNN with two pathways and seven deep layers (four convolutional layers, two fully connected layers, and one classification layer) to extract the ROI from the processed ulcer images. Each previously extracted modality is presented to the network as an input to one of its pathways. The aim of multiple pathways is to incorporate a larger amount of local contextual information from the ulcer images to extract the ROI. The last layer in the network is a 3D fully connected conditional random field (CRF) layer to enhance the performance of the network and remove false positives (FPs). CRF layer has the main advantage to overcome the limitations of other models, which is handling large neighborhoods in fast inference time. The DeepMedic network uses dense training on the segments that are extracted from the ulcer images to adapt class imbalance of this segmentation problem.

Fig. 5 Some examples of the output of the GS model. **a** The original ulcer images. **b** The equivalent output of the GS model



The 3D CNN network classifies each voxel in the processed image depending on the local and contextual information of its neighborhood to produce estimates for segmentation labels. It is done by multiple convolutions of the input of each pathway with some filters at the cascaded layers. Each layer consists of feature maps (FMs). Each FM is a group of neurons to detect a specific pattern in the previous layer. The voxel's receptive field is defined as voxels' neighborhood in the input modality that affects the activation of a neuron. Its size increases at each subsequent layer.

The result of the ROI extraction stage presents the external segmentation of the pressure ulcer image. In other words, the output of this stage introduces the ulcer area that is extracted from other background parts in the processed ulcer image. The output of the CNN is processed by morphological filters to enhance the resulting images. The used morphological filters are filling holes to remove discontinuity in the ulcer area, eliminating small areas, and applying closing operation, respectively. The result (\mathbf{m}_{ROI}) will be provided to the second 3D CNN network in the tissue segmentation stage as ROI input.

2.2 Tissue segmentation stage

The second stage of our proposed system is the tissue segmentation stage, which classifies the ulcer's tissues into granulation (red), necrotic eschar (black), and slough (yellow) tissues. Figure 4 illustrates the framework of this stage, which consists of four different modalities for the PU colored image. These modalities are fed as inputs to the 3D DeepMedic CNN network. The architecture of the second 3D CNN consists of eleven layers (eight convolutional

layers, two fully connected layers, and one classification layer) with four pathways. Each modality is supplied to one of the four pathways of the 3D CNN network. The ROI (\mathbf{m}_{ROI}), which is the output of the previous stage, and the GT (\mathbf{m}_{GT}) are also supplied to the 3D CNN network. The first two modalities are the same as in the previous stage. In addition, the architecture of the 3D DeepMedic CNN network is the same as used in the previous stage except that it has four pathways instead of two and eleven layers instead of seven layers.

The third modality (\mathbf{g}_L) is generated as an LCDG appearance model of the ulcer intensity image. The convolved intensity channel of the HSI images, which is generated in the last stage, is used to generate this modality. To build this modality, the collected empirical marginal probability distribution of voxel-wise intensities is approximated with a multimodal LCDG, and the latter is separated into the ulcer and background parts in unsupervised mode (see [22] for details).

LCDG creates a very close approximation of each distribution related to the mode with a linear combination of sign-alternate discrete Gaussian kernels. Each processed image is presented as a K-model image of dominant modes linked to the resulting image. In our case, K equals to three that represents the types of ulcer's tissues. The probability distributions of the tissues (F_s) are estimated and linked to each mode to segment the processed image by separating the modes:

$$F_s = \left(f_s(q) : q \in Q; \sum_{q \in Q} f_s(q) = 1 \right) \quad (3)$$

where q introduces the intensity levels that indicate the empirical marginal probability distribution of intensity levels for $mROI$. Therefore, the LCDG of the image is divided into sub-models linked to each dominant mode.

The discrete Gaussian (DG) is represented as the probability distribution (Ψ_θ) on Q where each probability ($\psi(q|\theta)$) relates to the cumulative Gaussian probability function ($\Phi_\theta(q)$) as follows:

$$\Psi_\theta = (\psi(q|\theta) : q \in Q) \quad (4)$$

$$\psi(q, \theta) = \begin{cases} \Phi_\theta(0.5) & \text{for } q=0, \\ \Phi_\theta(q+0.5) - \Phi_\theta(q-0.5) & \text{for } q=1, \dots, Q-2, \\ 1 - \Phi_\theta(q-0.5) & \text{for } q=Q-1. \end{cases} \quad (5)$$

where $\theta(\mu, \sigma^2)$, μ is the mean, and σ^2 is the variance. Finally, the LCDG with positive (C_p) and negative (C_n) components is calculated as follows:

$$p_{w,\Theta}(q) = \sum_{r=1}^{C_p} w_{p,r} \psi(q|\theta_{p,r}) - \sum_{l=1}^{C_n} w_{n,l} \psi(q|\theta_{n,l}) \quad (6)$$

$$\sum_{r=1}^{C_p} w_{p,r} - \sum_{l=1}^{C_n} w_{n,l} = 1 \quad (7)$$

where $C_p \geq K$ and its weights $w = [w_p, w_n]$ are nonnegative and satisfy Eq. (6). Finally, we aim to find a K-model probability that approximates the unknown marginal intensity level distribution, which can be calculated from:

$$g_L = \sum_{q \in Q} f(q) \log(p_{w,\theta}(q)) \quad (8)$$

Fig. 6 The step-by-step generation of LCDG modality for the colored ulcer image. **a** The original ulcer image. **b** The convolved intensity channel of the HSI image. **c** Estimated tissue classes by LCDG. **d** The output of the LCDG stage

where $f(q) = (|R|f_s(q) + 1)/(|R| + q)$ that represent its Bayesian estimate [23]. The model maximizes the expected log-likelihood of the statistically independent empirical data. Figure 6 shows the step-by-step generation of the LCDG modality for an ulcer image.

Finally, **the fourth modality (\mathbf{g}_{pr})** is the prior visual appearance model that uses the prior color information and the Euclidean distance to generate the color probability of the ulcer tissues. \mathbf{g}_{pr} is constructed by using three main steps. First, databases of the three classes of the ulcer tissues are generated by pixels' HSI values. These values are extracted from manually labeled images by ulcer experts. For each class $t \in \{1, 2, 3\}$, a database \mathbf{DB}_t is generated from aggregating the most repeated 100 HSI values of that class from the training labeled images. Second, the Euclidean distances ($\mathbf{D}_{r,t}$) are calculated between each pixel (\mathbf{r}) in the testing image (\mathbf{g}) and the three constructed tissue databases (\mathbf{DB}_t). Finally, all distances that are greater than or equal to a predefined threshold ($\mathbf{D}_{r,t} \geq \mathbf{T}$) are returned and used to calculate the prior probability for the current pixel:

$$P_{p,i} = \frac{N_t}{\sum_{t \in \{1,2,3\}} N_t} \quad (9)$$

where t is the tissue class (1 for necrotic eschar class, 2 for granulation class, and 3 slough class) and N is the total number of voxels that have a Euclidean distance greater than or equal the threshold. Figure 7 shows some examples of ulcer images with their equivalent prior visual appearance images.

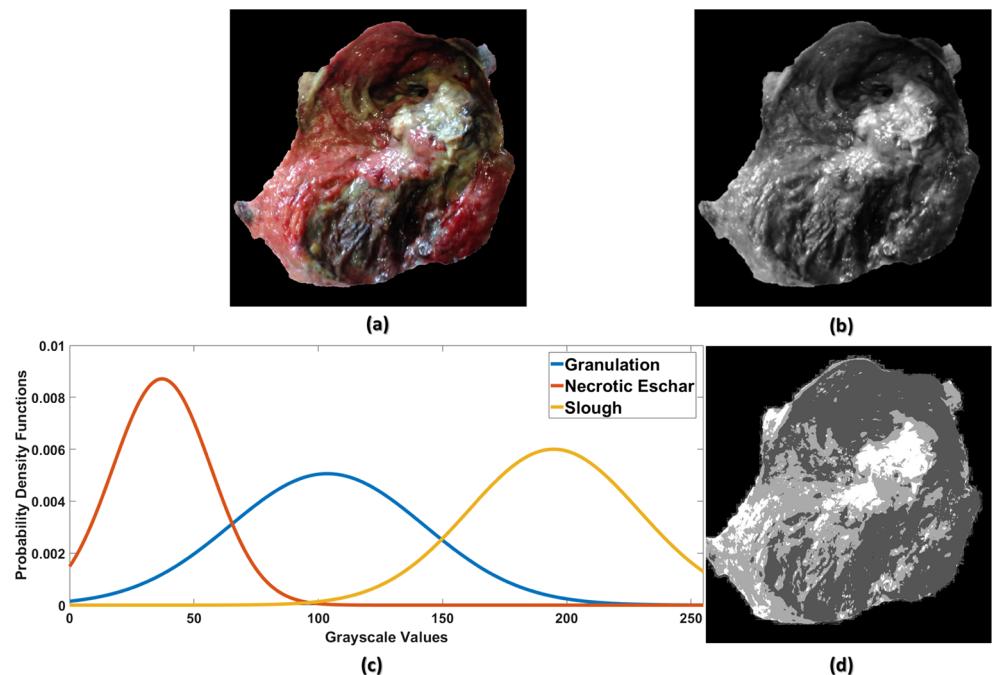
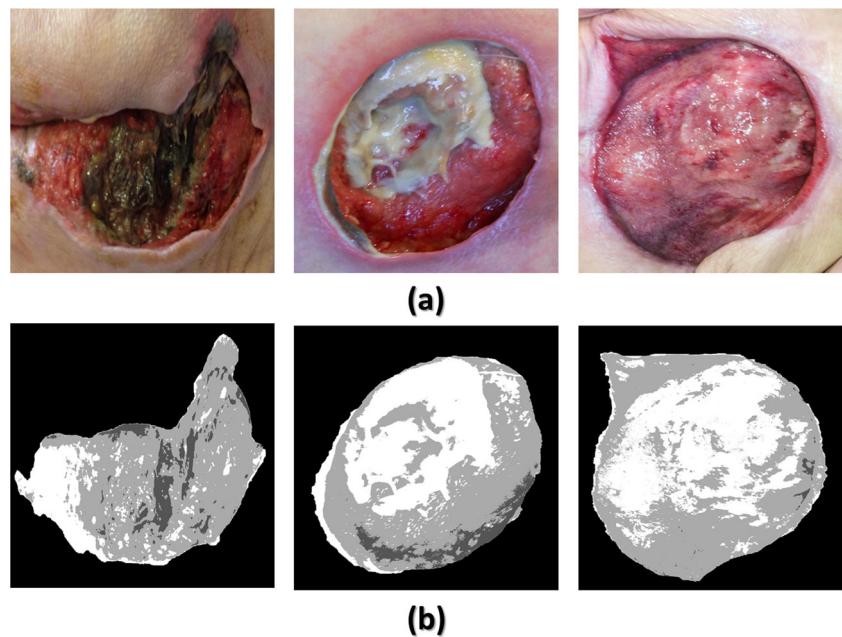


Fig. 7 Some examples of the output of the prior visual appearance model. **a** The original ulcer images. **b** The equivalent output of the g_{pr} model



After generating the four different modalities, they are supplied to the 3D CNN network as individual inputs to each pathway. The GT and ROI, which is generated from the first stage, are also supplied to the CNN network. Table 1 lists the main features of the used DeepMedic network.

3 Results

To test the proposed system, a data set of RGB colored images for PU is provided by IGURCO GESTIÓN S. L., which is a healthcare services company with many geriatric hospitals and nursing homes in the Basque Country (Spain). The data set contains 36 images with a resolution of 1024×1024 pixels, which are captured by using a regular digital camera. In addition, we got 157 images from Medetec wound database that have a resolution of 1024×731 pixels [24]. Three graduate students created

the GT for external boundaries and internal ulcer tissues. Finally, the images are refined, validated, and approved by three observers. We implemented our system on a Dell workstation with Intel Xeon CPU E5–2620 v4 at 2.10GHz (2 Quad Processors) and 64GB RAM. All algorithms are implemented by using Matlab, C++, and Python.

We evaluated the performance of the proposed system by using three common metrics, which are Dice similarity coefficient (DSC) [25], percentage area distance (PAD), and area under the curve (AUC). The DSC measures relevant correspondence between two areas regarding their true/false positive and negative parts. The PAD is the relative absolute ulcer difference between the results of the tested model and the GT segmentation. The AUC is the area under the Receiver Operating Characteristic (ROC) curve. It indicates the expectation that a uniformly drawn random positive is ranked before a uniformly drawn random negative. The lower the value of PAD and the higher the values of DSC

Table 1 The main characteristics of the 3D CNN for the tissue segmentation stage

Feature	Value
Number of layers	11 layers, which are eight convolutional layers (two 30×21^3 , two 40×17^3 , two 40×13^3 , and two 50×9^3), two fully connected layers (150×9^3), and one classification layer (2×9^3)
Residual connections	Add the input of layers 3, 5, and 7 to the outputs of layers 4, 6, 8, respectively.
Number of epochs in training	35 epochs with 20 sub-epochs for each one.
Learning rate	0.001
Momentum coefficient	0.6
Rho coefficient	0.9 with $\text{epsilon Rms} = 10^{-4}$
Regularization	$L_1 = 0.000001$ and $L_2 = 0.0001$

and AUC, the more accurate of our proposed segmentation system. DSC, PAD, and AUC are calculated by using Eqs. (10), (11), and (12), respectively [26].

$$DSC = \frac{200TP}{2TP + FP + FN}\% \quad (10)$$

$$PAD = 100 \frac{|S_a - GT_a|}{GT_a}\% \quad (11)$$

$$AUC = 0.5 \left(\frac{TP}{TP + FN} + \frac{TN}{TN + FP} \right) \quad (12)$$

where TP (true positive) is the total number of correctly classified pixels of the ulcer; TN (true negative) is the

number of the correctly classified pixels of the background; FP (false positive) is the total number of the misclassified pixels of the background; FN (false negative) is the total number of the misclassified pixels of the ulcer, S_a is the segmentation area, and GT_a is the ground truth area. To validate our results and prevent overfitting, we used two different cross-validation techniques. First, we applied the four-fold cross-validation. Second, we divided the dataset into 60% for training, 10% for validation, and 30% for testing.

Figure 8 shows some examples of PU images and their extracted ROI images. It illustrates the original RGB colored image, the GT, the output of the CNN network, and

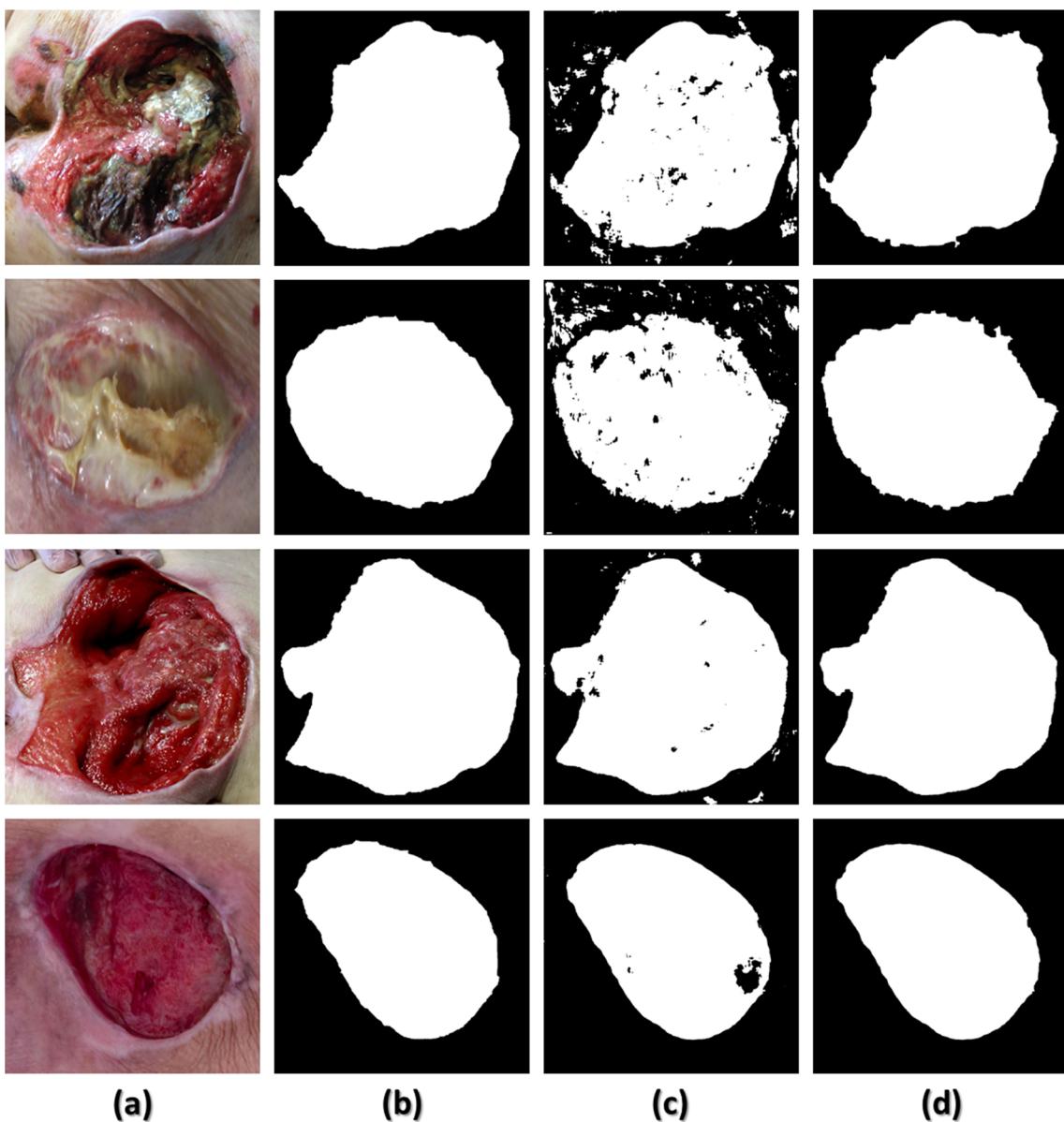


Fig. 8 Some examples of the output of the ROI extraction stage. **a** The original ulcer images. **b** The manually segmented ROI. **c** The output of the first 3D CNN. **d** The final output of the ROI extraction stage after the morphological filter

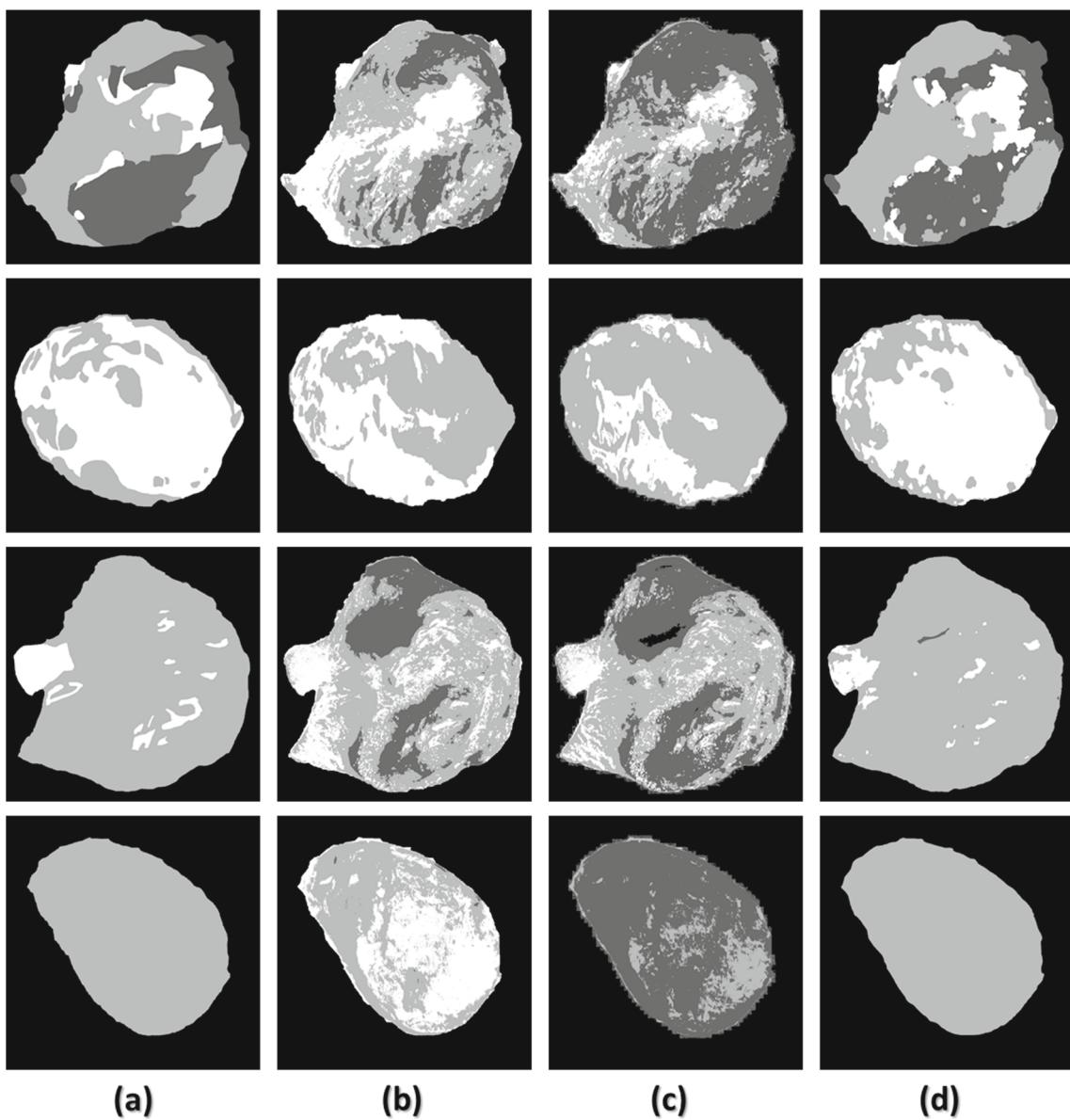


Fig. 9 Some examples of the output of the tissue classification stage. **a** The ground truth. **b** The output of the FCM. **c** The output of the LCDG. **d** The final output of our proposed system

the final output of the ROI extraction stage after applying morphological filters. To test the performance of the ROI extraction stage, we compared our result to the GT of the ROI by using three different metrics, which are DSC, PAD, and AUC. The ROI extraction stage achieved $DSC = 95 \pm 10\%$, $PAD = 8 \pm 12\%$, and $AUC = 95 \pm 8\%$. These results show that the first stage of our proposed system can detect the ulcer area automatically with high accuracy.

To measure the performance of our tissue segmentation system, we compared our proposed system with two other segmentation systems. The first is the segmentation output from the Fuzzy C-Means (FCM) algorithm [27]. The

second is the resulting segmentation from LCDG technique. Figure 9 shows some examples of the output of tissues segmentation systems. It illustrates the GT and the outputs of all tested systems including the results of our proposed system.

Table 2 The performance evaluation of the ROI extraction stage

Systems	DSC	PAD	AUC
Otsu segmentation	80 ± 20	17 ± 17	81 ± 17
The proposed system	95 ± 10	8 ± 12	95 ± 8

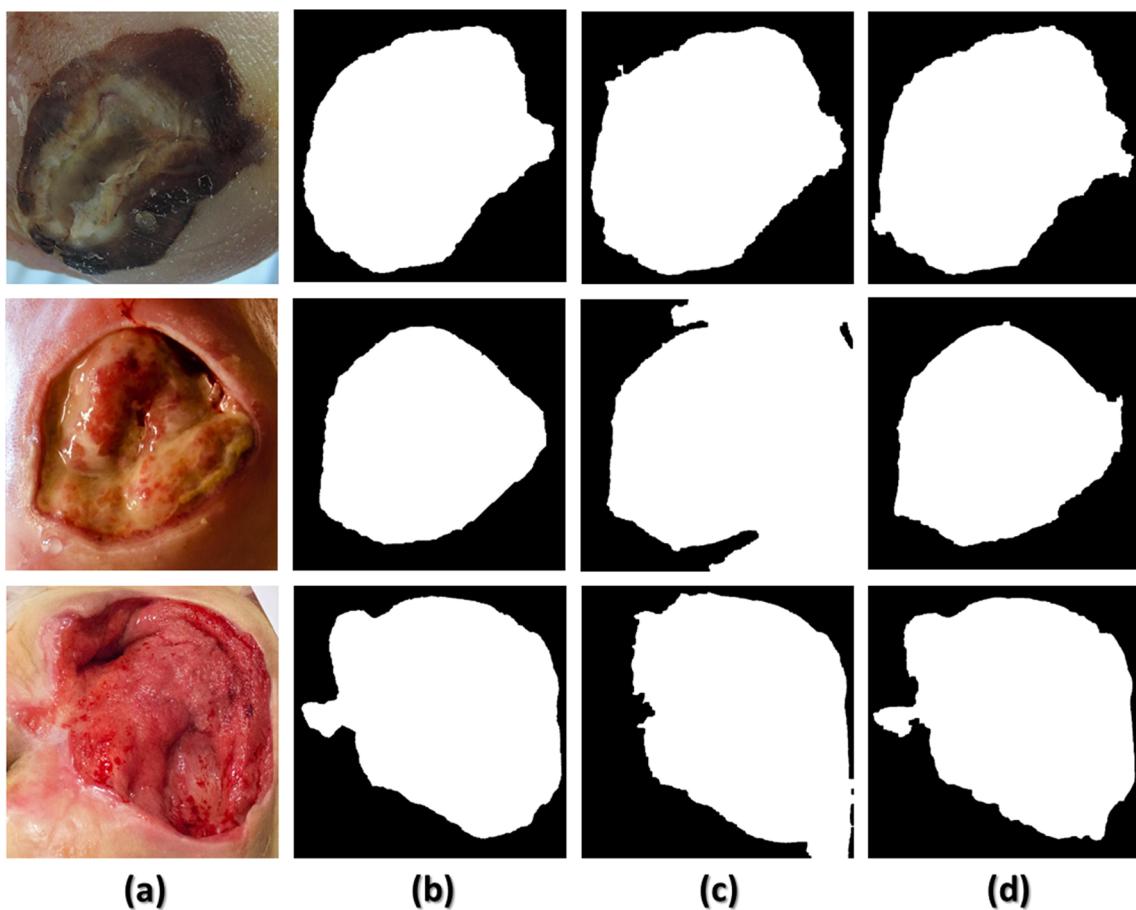


Fig. 10 The output of the ROI extraction stage by using Otsu segmentation and our proposed system. **a** The original ulcer images. **b** The manually segmented ROI. **c** The output of Otsu segmentation. **d** The output of our ROI extraction stage

4 Discussion

For the ROI extraction stage, we compared our results with Otsu segmentation technique [28]. Table 2 lists the results of the ROI extraction stage by using Otsu technique and our system. We applied the same morphological filters for both tested methods by using the same parameters. The results of

our system are obtained by dividing the data set into 60% for training, 10% for validation, and 30% for testing. Figure 10 shows some examples of PU images and their extracted ROI images by using Otsu technique and our proposed system.

For tissue segmentation, we compared our system with two other segmentation techniques, which are FCM and LCDG. Table 3 lists the experimental results of the testing

Table 3 The performance evaluation of three different systems for tissues segmentation in pressure ulcer images

Systems	Metrics	Necrotic eschar	Granulation	Slough
FCM segmentation	DSC	40 ± 37	69 ± 22	71 ± 10
	PAD	48 ± 41	35 ± 43	290 ± 102
	AUC	49 ± 29	77 ± 20	81 ± 8
LCDG segmentation	DSC	68 ± 12	70 ± 25	57 ± 30
	PAD	36 ± 21	59 ± 85	111 ± 49
	AUC	82 ± 8	80 ± 15	68 ± 21
The proposed system	DSC	92 ± 4	92 ± 6	91 ± 5
	PAD	7 ± 6	10 ± 18	22 ± 33
	AUC	93 ± 4	96 ± 3	95 ± 5

Table 4 The comparison between the proposed system using four different modalities with the 3D CNN networks with only one input modality

Tested modality	Metrics	Necrotic eschar	Granulation	Slough
HSI model	DSC	58 ± 6	67 ± 25	66 ± 27
	AUC	75 ± 7	82 ± 12	75 ± 27
GS model	DSC	21 ± 11	72 ± 26	69 ± 20
	AUC	54 ± 3	82 ± 14	78 ± 18
LCDG model	DSC	10 ± 3	80 ± 19	68 ± 24
	AUC	55 ± 2	90 ± 15	76 ± 27
Prior model	DSC	15 ± 5	79 ± 18	77 ± 18
	AUC	56 ± 3	88 ± 15	82 ± 22
The proposed system	DSC	92 ± 4	92 ± 6	91 ± 5
	AUC	93 ± 4	96 ± 3	95 ± 5

three different systems to evaluate the tissues segmentation in PU RGB images. We calculated the DSC, PAD, and AUC for necrotic eschar, granulation, and slough tissues. FCM segmentation achieved 60% for DSC, 124% for PAD, and 69% for AUC in average for all tissues. The high value of PAD indicates that there are tissues in some cases that are not classified correctly, especially in the slough tissue. LCDG segmentation shows better segmentation than FCM. It achieved 65% for DSC, 69% for PAD, and 77% for AUC in average for all tissues. Finally, our proposed system shows a high performance as compared with other systems. It achieved 92% for DSC, 13% for PAD, and 95% for AUC in average for all tissues. The results of PDA shows a high performance as compared with other systems, but it is still large in slough tissue. Similar to the ROI extraction stage, the results of our system are obtained by dividing the data set into 60% for training, 10% for validation, and 30% for testing. In addition to the last validation method, we validated our system by using four-fold cross-validation method. It achieved 91 ± 7% for DSC, 14 ± 18% for PAD, and 95 ± 4% for AUC in average for all tissues, which is consistent with the first validation method.

On the other hand, we compared the proposed tissue segmentation system using the four suggested modalities with four different 3D CNN networks. Each of these CNN networks is supplied with only one input modality, which are HSI image, GS model, LCDG model, and prior model. The aim of this comparison is to know the effect of fusing the four suggested modalities on the performance of the proposed system. Table 4 lists the experimental results of the testing five 3D CNN networks. Also, we calculated the DSC and AUC for necrotic eschar, granulation, and slough tissues. We noticed that the highest classification scores for the necrotic eschar tissue is from HSI model. On the other hand, the HSI model has the lowest scores for the other tissues. As proposed in our system, we used the four different modalities to have the highest classification score for the three various tissues.

As mentioned previously, three different experienced observers manually segmented the ulcer images. To validate the GT, the Bland-Altman analysis [29] is used to assess the degree of agreement between the GTs of the three observers. For good agreement, the mean difference between two GTs of two observers, which is called bias, is near zero. In addition, most of the data points should fall within 95% limits of agreement with the ±1.96 standard deviation (SD). We calculated this statistical analysis to compare the resulting ulcer area for each tissue type by two different observers. The Bland-Altman analysis confirms the robustness of the GTs of the three observers.

5 Conclusion

In this paper, we presented an automatic segmentation system to segment granulation, necrotic eschar, and slough tissues from pressure ulcer RGB images. The proposed system consists of two main stages that are ROI extraction and tissue classification. ROI extraction stage automatically extracts the ROI from the processed image to distinguish the PU area from other patient's skin and background objects. Tissue classification stage segments the resulting ROI to extract different ulcer tissues. Four different modalities of the processed images are supplied to a 3D CNN, which are the HSI image and its convolved version by using 3D Gaussian kernel, as well as its first-order current and prior visual appearance models using LCDG. The proposed system was trained and tested on 193 color PU images. The accuracy and robustness of the classifier were evaluated using the DSC, PAD, and AUC. The obtained preliminary results: DSC 92%, PAD 13%, and AUC 95% - are promising. In the future, we will extract some other modalities to increase the performance of the system. Also, we will increase the number of ulcer images to obtain more reliable results. Finally, we will make an assessment to the

PU depends on depth images to estimate the grade of the ulcer.

Acknowledgements The authors thank Prof. Dr. Georgy Gimel’farb, Department of Computer Science, University of Auckland, Auckland, New Zealand, for his help in revising the paper. In addition, the authors want to thank Sofia Zahia, Connor Burns, and Daniel Sierra-Sosa for their support in summarizing the related work and preparing the masks for the GT images.

Funding information The grants that have contributed with partial funding of the study are *IT* – 905 – 16 to eVIDA research group from the Basque Government, *JC2015* – 00305 Josè Castillejo Research Stay Grant from the Spanish Ministry, and *ACM2017.09* from the University of Deusto.

References



Begoña García-Zapirain is an associate professor at University of Deusto faculty. She published more than 35 papers in international scientific journals and presented more than 160 papers in international and national scientific conferences.



Ayman El-Baz is a professor, University Scholar, and Chair of the Bioengineering Department at the University of Louisville, KY. He has authored or co-authored more than 450 technical articles.



Mohammed Elmogy is an associate professor at Faculty of Computers and Information, Mansoura University, Egypt. He authored or co-authored over 120 publications in recognized international journals and conferences.



Adel S. Elmaghraby is a professor and chair of the Computer Engineering and Computer Science Department at the University of Louisville. He is a well-published author (over 250 publications), a public speaker, member of editorial boards, and technical reviewer.